THE GENERALIZED INVERSE OF A NONNEGATIVE MATRIX

R. J. PLEMMONS AND R. E. CLINE

Abstract. Necessary and sufficient conditions are given in order that a nonnegative matrix have a nonnegative Moore-Penrose generalized inverse.

1. Introduction. Let $A$ be an arbitrary $m \times n$ real matrix. Then the Moore-Penrose generalized inverse of $A$ is the unique $n \times m$ real matrix $A^+$ satisfying the equations

$$A = AA^+A, \quad A^+ = A^+AA^+, \quad (AA^+)^T = AA^+, \quad \text{and} \quad (A^+A)^T = A^+A.$$  

The properties and applications of $A^+$ are described in a number of papers including Penrose [7], [8], Ben-Israel and Charnes [1], Cline [2], and Greville [6]. The main value of the generalized inverse, both conceptually and practically, is that it provides a solution to the following least squares problem: Of all the vectors $x$ which minimize $\|b - Ax\|$, which has the smallest $\|x\|^2$? The solution is $x = A^+b$.

If $A$ is nonnegative (written $A \geq 0$), that is, if the components of $A$ are all nonnegative real numbers, then $A^+$ is not necessarily nonnegative. In particular, if $A \geq 0$ is square and nonsingular, then $A^+ = A^{-1} \geq 0$ if and only if $A$ is monomial, i.e., $A$ can be expressed as a product of a diagonal matrix and a permutation matrix, so that $A^{-1} = DA^T$ for some diagonal matrix $D$ with positive diagonal elements. The main purpose of this paper is to give necessary and sufficient conditions on $A \geq 0$ in order that $A^+ \geq 0$. Certain properties of such nonnegative matrices are then derived.

2. Results. In order to simplify the discussion to follow, it will be convenient to introduce a canonical form for a nonnegative symmetric
idempotent matrix. Flor [5] has shown that if \(E\) is any nonnegative idempotent matrix of rank \(r\), then there exists a permutation matrix \(P\) such that

\[
PEP^T = \begin{pmatrix}
J & JB & 0 & 0 \\
0 & 0 & 0 & 0 \\
AJ & AJB & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

where \(A\) and \(B\) are arbitrary nonnegative matrices of appropriate sizes and

\[
J = \begin{pmatrix}
J_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & J_r
\end{pmatrix}
\]

with each \(J_r\) a nonnegative idempotent matrix of rank 1. This gives the following lemma.

**Lemma 1.** Let \(E \succeq 0\) be a symmetric idempotent matrix of rank \(r\) with \(q\) nonzero rows. Then there exists integers \(\lambda_1, \cdots, \lambda_r\) and a permutation matrix \(P\) such that \(q = \lambda_1 + \cdots + \lambda_r\) and such that \(PEP^T\) has the form

(1) \[
PEP^T = \begin{pmatrix}
J_1 & 0 & & & \\
\vdots & \ddots & \vdots & & \\
0 & \cdots & J_r & 0 & \\
& & & \ddots & \\
& & & & \ddots & \\
0 & 0 & & & 0
\end{pmatrix}
\]

where each \(J_i\) is a \(\lambda_i \times \lambda_i\) positive idempotent matrix of rank 1.

The main result is given next. The theorem characterizes \(A \succeq 0\) so that \(A^+ \succeq 0\), and its proof indicates a method by which such an \(A^+\) can be constructed readily.

**Theorem 1.** Let \(A\) be an \(m \times n\) nonnegative matrix of rank \(r\). Then the following statements are equivalent.

(i) \(A^+\) is nonnegative.
(ii) There exists a permutation matrix $P$ such that $PA$ has the form
\[
PA = \begin{pmatrix}
B_1 \\
\cdot \\
\cdot \\
B_r \\
0
\end{pmatrix}
\]
where each $B_i$ has rank 1 and where the rows of $B_i$ are orthogonal to the rows of $B_j$ whenever $i \neq j$.

(iii) $A^+ = DA^T$ for some diagonal matrix $D$ with positive diagonal elements.

PROOF. Suppose (i) holds so that $A, A^+ \succeq 0$. Since $E = AA^+$ is a symmetric idempotent, there exists a permutation matrix $P$ so that $K = PEPT$ has the form (1). Let $B = PA$. Then $B^+ = A^+P^T$, $BB^+ = K$, $KB = B$, and $B^+K = B^+$. Now $B$ can be partitioned into the form (2), where $r$ is the rank of $A$ and where each $B_i$, $1 \leq i \leq r$, is a $\lambda_i \times n$ matrix with no zero rows, since $A$ and $B$ have the same number of nonzero rows. It remains to show each $B_i$ has rank 1 and $B_iB_j^T = 0$, for $1 \leq i \neq j \leq r$. Let $C = B^+$. Then $C$ can be partitioned into the form
\[
C = (C_1, \cdots, C_r, 0)
\]
where, for $1 \leq i \leq r$, $C_i$ is an $n \times \lambda_i$ matrix with no zero columns. Moreover, since $CB$ is symmetric, a column of $B$ is nonzero if and only if the corresponding row of $C$ is nonzero. Now $KB = B$ implies that $J_iB_i = B_i$, so that $B_i$ has rank 1, for $1 \leq i \leq r$. It remains to show that the rows of $B_i$ are orthogonal to the rows of $B_j$ for $i \neq j$. Since $BC = K$ has the form (1),
\[
B_iC_j = J_i, \quad \text{if} \quad i = j, \quad \text{and}
\]
\[
= 0, \quad \text{if} \quad i \neq j,
\]
for $1 \leq i, j \leq r$. Suppose the $l$th column of $B_i$ is nonzero. Then $B_iC_k = 0$ for $k \neq i$ implies that the $l$th row of $C_k$ is zero. However, since the $l$th row of $C$ is nonzero, the $l$th row of $C_i$ is nonzero. In this case, the $l$th column of $B_k$ is zero for all $k \neq i$, since $B_kC_i = 0$. Thus
\[
B_iB_j^T = 0 \quad \text{for all} \quad 1 \leq i \neq j \leq r,
\]
and (ii) is established.

---

2 Note that the zero block may not be present.
Now assuming (ii) holds, let \( B = PA \) have the form (2). Then for
\[ 1 \leq i \leq r, \]
there exist column vectors \( x_i, y_i \) such that \( B_i = x_i y_i^T \). Furthermore, \( B_i^+ \) is the nonnegative matrix
\[ B_i^+ = (\|x_i\|_2^2 \|y_i\|_2^2)^{-1}B_i^T \]
and moreover \( B^+ = (B_1^+, \cdots, B_r^+, 0) \), since \( B_i B_j^T = 0 \) for \( i \neq j \). In particular then, \( B^+ = DB^T \) where \( D \) is a diagonal matrix with positive diagonal elements and thus \( A^+ = DA^T \), yielding (iii).

Clearly (iii) implies (i) so the proof is complete.

The next theorem considers doubly stochastic matrices, that is, square matrices \( A \geq 0 \) whose row sums and column sums are 1. The matrix \( A \geq 0 \) is said to be diagonally equivalent to a doubly stochastic matrix if there exist diagonal matrices \( D_1 \) and \( D_2 \) such that \( D_1 A D_2 \) is doubly stochastic. Classes of nonnegative matrices with this property have been the subject of several recent papers (for example, see Djoković [4]). Part of the following theorem identifies another such class.

**Theorem 2.** Let \( A \geq 0 \) be square with no zero rows or columns. If \( A^+ \geq 0 \) then \( A \) is diagonally equivalent to a doubly stochastic matrix. Moreover, if \( A \) is doubly stochastic then \( A^+ \) is doubly stochastic if and only if the equation \( A = AXA \) has a doubly stochastic solution, in which case \( A^+ = A^T \).

**Proof.** The first statement follows since there exist permutation matrices \( P \) and \( Q \) such that
\[
P A Q = \begin{pmatrix}
B_1 & & 0 \\
& \ddots & \\
0 & & B_r
\end{pmatrix}
\]
where each \( B_i \) is a positive square matrix.

For the second statement note that a doubly stochastic idempotent matrix \( E \) is necessarily symmetric; for in particular, there exists a permutation matrix \( P \) such that \( PEP^T \) has the form (1), where each row and column is nonzero and where each \( J_i \) is a positive, idempotent doubly stochastic matrix of rank 1. Then each entry of \( J_i \) is \( 1/\lambda_i \) so that \( PEP^T \) and, accordingly, \( E \) are symmetric matrices. This means that \( A^+ \) is the only possible doubly stochastic solution to the equations \( A = AXA \) and \( Y = YAY \), since \( AY \) and \( YA \) are symmetric and \( A^+ \) is unique. Thus \( A^+ \) is doubly stochastic if and only if \( A = AXA \) has a doubly stochastic solution, in which case \( A^+ = XAX \), and so \( A^+ = A^T \) by Theorem 1.

The final result determines the singular values of \( A \) (i.e., the positive square roots of the nonzero eigenvalues of \( A^T A \)) whenever \( A^+ \geq 0 \).
Theorem 3. Let \( A \preceq 0 \) be an \( m \times n \) real matrix with \( A^+ \succeq 0 \) and let \( PA \) have the form (2). Let \( \{x_i, y_i\}_{i=1}^r \) be column vectors so that \( B_i = x_i y_i^T \) for \( 1 \leq i \leq r \). Then the singular values of \( A \) are the numbers \( \|x_i\| \cdot \|y_i\| \).

Proof. The eigenvalues of \( AA^T \) are the eigenvalues of \( BB^T \). But these are the eigenvalues of the matrices \( B_i B_i^T \) for \( 1 \leq i \leq r \), that is, the numbers \( \|x_i\|^2 \cdot \|y_i\|^2 \).

References